NORTH-HOLLAND

A Relative Perturbation Bound for Positive Definite Matrices

Roy Mathias*
Department of Mathematics
College of William and Mary
Williamsburg, Virginia 23187

and

Krešimir Veselić†
Lehrgebiet Mathematische Physik
Fernuniversität Hagen
Postfach 940
D-58084 Hagen, Germany

Submitted by Ludwig Elsner

ABSTRACT

We give a sharp estimate for the eigenvectors of a positive definite Hermitian matrix under a floating-point perturbation. The proof is elementary. © 1998 Elsevier Science Inc.

Recently there have been a number of papers on eigenvector perturbation bounds that involve a perturbation of the matrix which is small in some relative sense, including the typical rounding errors in matrix elements [1, 2, 10, 8, 4, 5]. The proofs are mostly complicated, and all of them involve the notion of the *relative gap* between the eigenvalues, i.e. a relative distance

*E-mail: mathias@math.wm.edu. Research supported in part by a grant from the National Science Foundation and a Summer Research Grant for the College of William and Mary.
†E-mail: kresimir.veselic@fernuni-hagen.de.

655 Avenue of the Americas, New York, NY 10010

0024-3795/98/$19.00
PII S0024-3795(97)00248-6

from the unperturbed eigenvalue to the rest of the spectrum. Several such relative gaps are in use. In any such estimate it is only the nearest eigenvalue that matters; one does not care about distant eigenvalues and their influence. Our bounds control primarily the angles between the perturbed and the unperturbed eigenvectors (standard bounds with relative gaps may be derived from them any time). In particular, the distant eigenvalues naturally damp out the perturbation of the corresponding components of the eigenvector. The bounds are asymptotically sharp, i.e., for small perturbations they reach the first term of the perturbation theory. Our proof is simple (of all works cited above, [4, 5] are the closest to ours)—the only technical tool is the square root of a positive definite matrix. The simplicity of our proof may make it useful in a classroom.

THEOREM 1. *Let $H = U\Lambda U^*$ and $\tilde{H} = H + \delta H = \tilde{U}\tilde{\Lambda}\tilde{U}^*$ be positive definite. Assume that U and $\tilde{U}$ are unitary and that Λ and $\tilde{\Lambda}$ are diagonal. Let $S = U^*\tilde{U}$, and assume*

$$\eta = \|H^{-1/2}\,\delta H\, H^{-1/2}\| < 1. \tag{1}$$

Then for any j and for any set $\mathcal{S}$ not containing j we have

$$\left(\sum_{i\in\mathcal{S}} |s_{ij}|^2\right)^{1/2} \leqslant \max_{i\in\mathcal{S}} \frac{\lambda_i^{1/2}\tilde{\lambda}_j^{1/2}}{|\lambda_i + \tilde{\lambda}_j|}\frac{\eta}{\sqrt{1-\eta}} \tag{2}$$

and, in particular,

$$|s_{ij}| \leqslant \frac{\lambda_i^{1/2}\tilde{\lambda}_j^{1/2}}{|\lambda_i - \tilde{\lambda}_j|}\frac{\eta}{\sqrt{1-\eta}}. \tag{3}$$

This theorem implies a normwise bound even in the indefinite case.* Take a Hermitian matrix H and a perturbation δH. Consider $H + tI$ and $H + tI + \delta H$. These two matrices are positive definite for t sufficiently large, and have the same eigenvectors as H and $H + \delta H$ respectively. Now as $t \to \infty$,

$$\eta\lambda_i^{1/2}(H + tI)\lambda_j(H + \delta H + tI) \to \|\Delta H\|,$$
$$\eta\lambda_{\mathrm{i}}^{1/2}(\mathrm{I} + \mathrm{tI})\lambda_{\mathrm{j}}(\mathrm{H} + \delta\mathrm{H} + \mathrm{tI}) \to \|\Delta\mathrm{H}\|,$$

*This fact, together with its proof, is due to the referee.

where $\eta = \|(H + tI)^{1/2}\, \delta H\, (H + tI)^{-1/2}\|$. The bound (2) implies

$$\left(\sum_{i \in \mathcal{S}} |s_{ij}|^2\right)^{1/2} \leqslant \max_{i \in \mathcal{S}} \frac{\|\delta H\|}{|\lambda_i^{1/2} - \tilde{\lambda}_j^{1/2}|}.$$

Proof. Let $X = \Lambda^{-1/2} U^* \,\delta H\, U \Lambda^{-1/2}$. Then

$$\begin{aligned}\|X\| &= \|\Lambda^{-1/2} U^* \,\delta H\, U \Lambda^{-1/2}\| = \|U\Lambda^{-1/2} U^* \,\delta H\, U \Lambda^{-1/2} U^*\| \\ &= \|H^{-1/2}\, \delta H\, H^{-1/2}\| = \eta.\end{aligned}$$

Also

$$\begin{aligned}\tilde{U}\tilde{\Lambda}\tilde{U}^* &= \tilde{H} = H + \delta H = U\Lambda^{1/2}(I + \Lambda^{-1/2} U^* \,\delta H\, U\Lambda^{-1/2})\Lambda^{1/2} U^* \\ &= U\Lambda^{1/2}(I + X)\Lambda^{1/2} U^*. \qquad (4)\end{aligned}$$

Now multiplying on the left by U^* and on the right by $\tilde{U}$ gives

$$S\tilde{\Lambda} = \Lambda S + \Lambda^{1/2} X \Lambda^{1/2} S.$$

Taking the ij element of this identity, we obtain

$$s_{ij} = \frac{\lambda_i^{1/2}}{\tilde{\lambda}_j - \lambda_i} (X\Lambda^{1/2} S)_{ij}. \qquad (5)$$

Thus, we need a bound on $X\Lambda^{1/2}S$. The equation (4) can be written as

$$I = (\tilde{\Lambda}^{-1/2} S^* \Lambda^{1/2})(I + X)(\Lambda^{1/2} S \tilde{\Lambda}^{-1/2}),$$

and hence

$$\|\Lambda^{1/2} S \tilde{\Lambda}^{-1/2}\| = \left\|(I + X)^{-1}\right\|^{1/2} \leqslant (1 - \eta)^{-1/2}.$$

Let $Y_{\cdot j}$ denote the jth column of the matrix Y. Then

$$\begin{aligned}\left\|(X\Lambda^{1/2}S)_{\cdot j}\right\| &= \left\|(X\Lambda^{1/2} S \tilde{\Lambda}^{-1/2} \tilde{\Lambda}^{1/2})_{\cdot j}\right\| = \tilde{\lambda}_j^{1/2} \left\|(X\Lambda^{1/2} S \tilde{\Lambda}^{-1/2})_{\cdot j}\right\| \\ &\leqslant \tilde{\lambda}_j^{1/2} \|X\|\, \|\Lambda^{1/2} S \tilde{\Lambda}^{-1/2}\| \leqslant \tilde{\lambda}_j^{1/2} \eta (1 - \eta)^{-1/2}.\end{aligned}$$

Thus,

$$
\begin{aligned}
\sum_{i \in \mathscr{S}} |s_{ij}|^2 &= \sum_{i \in \mathscr{S}} \left(\frac{\lambda_i^{1/2}}{\lambda_i - \tilde{\lambda}_j} \right)^2 \left| (X \Lambda^{1/2} S)_{ij} \right|^2 \\
&\leqslant \max_{i \in \mathscr{S}} \left(\frac{\lambda_i^{1/2}}{\lambda_i - \tilde{\lambda}_j} \right)^2 \sum_{i \in \mathscr{S}} \left| (X \Lambda^{1/2} S)_{ij} \right|^2 \\
&\leqslant \max_{i \in \mathscr{S}} \left(\frac{\lambda_i^{1/2}}{\lambda_i - \tilde{\lambda}_j} \right)^2 \left\| (X \Lambda^{1/2} S)_{\cdot j} \right\|^2 \\
&\leqslant \max_{i \in \mathscr{S}} \left(\frac{\lambda_i^{1/2}}{\lambda_i - \tilde{\lambda}_j} \right)^2 \tilde{\lambda}_j \eta^2 (1 - \eta)^{-1}.
\end{aligned}
$$

Taking square roots yields (2), which then implies (3). ■

The last inequality can be used to obtain a bound for the perturbation of the eigenvectors. We can obviously always choose $\tilde{U}$ such that $s_{jj} \geqslant 0$ for all j. Using $s_{jj} \geqslant 0$, we obtain

$$
\begin{aligned}
&\|\tilde{U}_{\cdot j} - U_{\cdot j}\| \\
&\quad = \sqrt{2}\sqrt{1 - s_{jj}} \\
&\quad = \sqrt{2}\sqrt{1 - \sqrt{1 - |s_{1j}|^2 - \cdots - |s_{j-1,j}|^2 - |s_{j+1,j}|^2 - \cdots - |s_{nj}|^2}}\ .
\end{aligned}
$$

Then by (2)

$$
\begin{aligned}
\|\tilde{U}_{\cdot j} - U_{\cdot j}\| &\leqslant \sqrt{2}\sqrt{1 - \sqrt{1 - \frac{\eta^2}{1-\eta} \max_{i(\neq j)} \frac{\lambda_i \tilde{\lambda}_j}{\left(\lambda_i - \tilde{\lambda}_j\right)^2}}} \\
&\leqslant \frac{\eta\sqrt{2}}{\sqrt{1-\eta}} \max_{i(\leqslant j)} \frac{\lambda_i^{1/2}\tilde{\lambda}_j^{1/2}}{|\lambda_i - \tilde{\lambda}_j|}
\end{aligned}
$$

Different choices of $\mathscr{S}$ would allow estimates for invariant subspaces.

The bounds (2, 3) involve $\tilde{\lambda}_j$ and λ_j in a symmetric way. One can write the bounds entirely in terms of the eigenvalues of H by using the fact [2, 6][1]

$$\lambda_j(1 - \eta) \leqslant \tilde{\lambda}_j \leqslant \lambda_j(1 + \eta). \tag{6}$$

Since the proof is also very simple, we repeat it here for convenience. Note that (1) implies $|x^* \delta H x| \leqslant \eta x^* H x$ for any vector x. Thus,

$$(1 - \eta) x^* H x \leqslant x^*(H + \delta H) x \leqslant (1 + \eta) x^* H x,$$

and (6) follows from the monotonicity property for the eigenvalues.[2] If η is small, then the right-hand side in (3) reads

$$\frac{\lambda_i^{1/2} \lambda_j^{1/2}}{|\lambda_i - \lambda_j|} \eta + \text{higher terms}. \tag{7}$$

On the other hand, for a simple eigenvalue λ_i the perturbation theory gives

$$u_i^* \tilde{u}_j = \frac{u_i^* \, \delta H u_j}{\lambda_i - \lambda_j} + \text{higher terms}.$$

Taking here e.g. $\delta H = \eta_0 \lambda_i^{1/2} \lambda_j^{1/2}(u_i u_j^* + u_j u_i^*)$, $\eta_0 < 1$, we obtain

$$u_i^* \tilde{u}_j = \frac{\lambda_i^{1/2} \lambda_j^{1/2}}{\lambda_i - \lambda_j} \eta_0 + \text{higher terms}.$$

Since here $\eta = \eta_0$, we see that our bound is asymptotically sharp.

Since the space dimension n does not enter the main estimates (2, 3), they will hold for corresponding perturbations of Hilbert-space operators with compact resolvent as well.

---

[1] Of course, here we assume that Λ and $\tilde{\Lambda}$ are equally ordered.

[2] The monotonicity is, of course, a consequence of the minimax theorem. Most perturbation estimates for the eigenvalues of Hermitian matrices, however, follow more directly from the monotonicity, which, in turn, is a fact much easier to keep in mind and memorize. This may be of relevance in teaching this matter.

Relative bounds in computational practice are seldom actually found in the form (1). Most common is the componentwise bound

$$|\delta H_{ij}| \leqslant \varepsilon |H_{ij}|, \tag{8}$$

where ε is e.g. the machine precision. It can be shown (see e.g. [10]) that (8) implies (1) with

$$\eta \leqslant \| |A| \| \|A^{-1}\| \varepsilon \leqslant \sqrt{n}\, \kappa(A) \varepsilon, \tag{9}$$

where $\|\cdot\|$ is the spectral norm, D is any diagonal matrix, and $H = DAD$. One chooses D so as to make the condition number $\kappa(A)$ as small as possible. As was shown in [9], a nearly optimal (within a factor $\sqrt{n}$) choice is $D = \sqrt{\operatorname{diag}(H)}$. The factor $\sqrt{n}$ in (9) cannot be dispensed with [7].

For the convenience of the reader we will derive the first inequality in (9):

$$|x^* \delta H x| \leqslant |x|^T |\delta H| \, |x| \leqslant \varepsilon |x|^T |A| D |x| \leqslant \varepsilon \| |A| \| x^* D^2 x$$

$$\leqslant \varepsilon \| |A| \| \, \|A^{-1}\| x^* H x.$$

This implies (1) with η from (9).

Although the same type of eigenvector estimates can be expected for indefinite Hermitian matrices as well (cf. e.g. [3, 10]), we have not been able to extend our simple method of proof to this case as yet.

REFERENCES

1 J. Barlow and J. Demmel, Computing accurate eigensystems of scaled diagonally dominant matrices, *SIAM J. Numer. Anal.* 27:762–791 (1988).

2 J. Demmel and K. Veselić, Jacobi's method is more accurate than QR, *SIAM J. Matrix Anal. Appl.* 13:1204–1245 (1992).

3 S. C. Eisenstat and I. C. Ipsen, Relative perturbation techniques for eigenspaces and singular vector subspaces, in *Proceedings of the Fifth SIAM Conference on Applied Linear Algebra*, 1994, pp. 62–66.

4 S. C. Eisenstat and I. C. Ipsen, Relative perturbation techniques for singular value problems. *SIAM J. Numer. Anal.* 32(6):1972–1988 (1995).

5 R.-C. Li, Relative perturbation theory: (II) Eigenspace and singular subspace variations, Manuscript.

6 R. Mathias, Spectral perturbation bounds for graded positive definite matrices, *SIAM J. Matrix Anal. Appl*, to appear.

7 R. Mathias, A perturbation bound for eigenvectors and for the matrix square root, *SIAM J. Matrix Anal. Appl.*, to appear.

8 I. Slapničar and K. Veselić, Perturbations of the eigenprojections of factorized Hermitian matrices, *Linear Algebra Appl.* 218:273–280 (1995).

9 A. van der Sluis, Condition numbers and equilibrations of matrices, *Numer. Math.* 195:14–23 (1993).

10 K. Veselić and I. Slapničar, Floating-point perturbations of Hermitian matrices, *Linear Algebra Appl.* 195:81–116 (1993).

Received 3 June 1996; final manuscript accepted 3 February 1997